

Ins
B1

B1

Voice-controlled telecommunication terminal

5 The present invention relates to a method for controlling a telecommunication terminal by means of voice, as presented in the preamble of the claim 1, and a voice-controlled telecommunication terminal according to the method.

10 When a mobile phone is used in a car, a hands-free mode is often required, wherein the car has a hands-free equipment for the mobile phone, comprising a separate loudspeaker and a microphone. Thus, the speaker can use both hands for driving during the call. The advantages of the hands-free mode are comfort in use and improved safety. To increase comfort in use, the hands-free mode is used also in offices as a desktop hands-free installation.

15 The convenience of hands-free mode is decreased by the fact that for making a call the driver has to dial the telephone number by pressing the keys of the phone. This impairs traffic safety, because the look of the driver is attached to the phone. To facilitate the dialling of numbers, shortcut functions have been designed to phones, wherein names and numbers of persons have been stored into the memory of the phone. The shortcut memory can be scrolled through, wherein it is advantageous to show on the display device of the phone an identifier corresponding to each telephone number, such as the name of the respective person. If needed, it is also possible to show the phone number corresponding to the identifier. The memory can be scrolled forwards and backwards, and when the desired identifier appears on the display device, the dialling the phone number can be started, for example by pressing a call key. However, the shortcut function does not entirely eliminate the need to press the keys when calling.

20
25
30

35 Various methods based on voice recognition for telecommunication terminals, such as mobile phones and wireline telephones, have been developed, particularly for dialling a phone number without pressing the keys. In such methods, the desired phone number can be dialled usually in a manner that the caller pronounces the phone number or an identifier related to the phone number, such as the name of the person.

00918386-002657

DL
826-48

The phone number corresponding to the identifier has been stored advantageously to the shortcut memory.

5 Some known telecommunication terminals and methods based on voice recognition have been described in the patent publications US-4,644,107, US-4,853,953, US-4,928,302, US-5,182,765 and US-5,222,121.

10 Prior art control and calling methods of a telecommunication terminal using voice recognition are mainly based on the fact that a distinguishing voice pattern has been stored for each command and phone number. Thus, the command or identifier has to be given in a form as identical with the stored form as possible. Thus, the caller has to remember in which form e.g. the name "Matthew Herbert Williams" was stored;
15 was it stored exactly in this form, or in form "Matthew Williams", "Williams Matthew", or "Williams Matthew Herbert".

US-patent 5,222,121 discloses a voice-recognition dialling device arranged in connection with a telephone mounted on a vehicle or ^{the} a like.
20 Into the memory of the dialling device, voice patterns corresponding to the commands and telephone numbers, such as words "RECALL MEMORY", "SEND" and "VERIFY", are stored. Voice patterns are preferably stored already when the dialling device is manufactured. The dialling unit can also be implemented in a manner that the user teaches
25 the unit also the commands and numbers. The dialling device includes a loudspeaker and/or a display device, wherein the user is given instructions in form of voice signals and/or text. The call is initiated by pronouncing the command "RECALL MEMORY", wherein the dialling device requests the user to pronounce the identifier of the desired telephone number. After the identifier has been pronounced, the device
30 compares the identifiers stored into the memory and after finding an identifier that most resembles the pronounced identifier, it gives a voice signal. The user may then give the device a call command "SEND", or a command "VERIFY" if the user wishes to check that the number is correct. In this case, the dialling device informs the chosen identifier, for example in a sound signal. If the chosen identifier is correct, a connection is created by using a call command. If the chosen identifier is
35 incorrect, the user can scroll through the other alternatives by using a

03913366.032607

24
8.26.48

command "NEXT ONE". However, the identifiers have to be given in the same form as they have been stored, which increases the possibility of false choices.

- 5 US-patent 4,928,302 presents another dialling device for calling a desired telephone number by using voice commands. In this device, the telephone numbers can be classified for example according to the initial part of the name. The search can thus be implemented by pronouncing for example the surname "Williams", wherein the device searches all
- 10 the names having "Will" in their initial part, such as "Williams", "Williamson" and "Willis". In the next phase the desired name can be chosen from the list formed by the device, which is thus in this phase briefer than the list of all the names stored in the memory. Even this device has the disadvantage that the user has to remember the form
- 15 the name was stored, that is, "Williams Matthew", "Matthew Williams", "Williams Matthew Herbert" or "Matthew Herbert Williams",

- 20 The purpose of the present invention is to eliminate the above mentioned disadvantages to a great extent and to provide a device and method for controlling a telecommunication terminal by means of voice command, particularly for choosing a telephone number from a group of stored telephone numbers. The invention is based on the idea that the identifier can comprise more than one sub-identifier, i.e. word, wherein in the search phase the identifier can be dictated according to combination of any sub-identifiers. The method of the invention is characterized in what is said in the characterizing portion of the appended claim 1. The voice-controlled unit of the invention is characterized in what is said in the characterizing portion of the appended claim 3.
- 25
- 30 The present invention provides significant advantages over prior art voice-control methods and voice-controlled devices.

- 35 In the method according to the invention the identifier related to a telephone number can be composed of one or several sub-identifiers stored into the memory of the device. However, it is not required in the calling phase to pronounce the sub-identifiers in the exact order as they were stored, but any combination or partial combination of sub-identifiers can be used. It is not even necessary to pronounce all the sub-

00916366-002697

8-25-98

Ins
B2

B2

identifiers provided that the telephone number to be chosen is identified by the group of the pronounced sub-identifiers. In some cases the identifier can be identified by pronouncing just one sub-identifier.

- 5 A method in accordance with a second advantageous embodiment of the invention provides the option to pronounce sub-identifiers not present in the group of sub-identifiers stored in the memory, that is the word list, when the telephone number is chosen. The voice-recognition advantageously ignores these sub-identifiers and performs the selection based on sub-identifiers present in the word list.

10 In the following, the invention is described in more detail with reference to the accompanying drawing, where

- 15 Fig. 1 shows a reduced block diagram of one advantageous dialling device according to the invention,
- Fig. 2 shows a reduced flow chart of storing of an identifier into the memory of the device, and
- 20 Fig. 3 shows a reduced flow chart of a situation in which a telephone number is dialled in accordance with one advantageous embodiment of the invention.

- 25 A voice-controlled telecommunication terminal 1 according to an advantageous embodiment of the invention as shown in Fig. 1 is for example a mobile station, such as a GSM mobile phone, or a fixed wire-line telephone. Fig. 1 shows only those blocks which are the most essential for understanding the invention. A voice-control unit 2 comprises
- 30 advantageously a voice-recognition means 3, a voice pattern memory 4, a controller unit 5, read-only memory 6, random access memory 7, speech synthesiser 8 and an interface 9. Voice control can be given e.g. by means of a microphone 10a of the telecommunication terminal 1 or by means of a microphone 10b of a hands-free equipment 17. The
- 35 instructions and notices to the user can be given e.g. by means of sound signals created by a speech synthesiser 8 either through a loudspeaker 11a belonging to the telecommunication terminal 1 or through a loudspeaker 11b of the hands-free equipment. The voice-control

00910386-002697

84
8.26.98

5

30

35

collected when calling. The user can give the telephone number also by keying in the corresponding numerals. The teach mode of the voice commands is terminated advantageously by keying again the voice-store key A or through the menu function of the telecommunication terminal.

5
10
15
20
25
30
35

In the phase when the user wishes to store the identifier of the telephone number, the voice-control unit 2 is set to a mode in which the voice-control unit can expect to receive identifiers, which can be composed of one or several sub-identifiers. This function mode is described in the following with reference to the flow chart of Fig. 2. Changing over to the store-identifier mode (block 201) is implemented advantageously by keying the voice-store key A or through the menu facility, as presented earlier in connection with command storing. The voice-recognition unit 2 creates advantageously a message "Pronounce the identifier" (block 202), wherein the user starts pronouncing the sub-identifiers of the identifier. Thus, the identifier can comprise one or several sub-identifiers, for example "Williams", "Matthew", "Herbert". A short pause is kept between each sub-identifier, wherein the voice-recognition unit 2 is able to separate the sub-identifiers from each other. Each pronounced sub-identifier is stored into the voice-equivalent memory 4 (block 203). The voice-control unit 2 can additionally create a short sound signal (e.g. a bleep) after each pronounced sub-identifier as a sign that the sub-identifier is stored. Subsequently, after all the sub-identifiers have been pronounced (block 204), the user is requested to give the telephone number related to the identifier (block 205), e.g. by pronouncing the numbers or by keying. After the number has been given, the voice-control unit 2 stores the telephone number e.g. to the random access memory 7 (block 206) and creates references of the sub-identifiers to the telephone number (block 207). Subsequently, the user is asked whether any other identifiers and telephone numbers are to be stored (blocks 209, 210). In case the user wishes to continue the storing, the function moves back to the block 202 until identifiers are no longer given (block 211).

Division into sub-identifiers can also be implemented in a manner that the user divides the identifier into sub-identifiers and separates the sub-identifiers e.g. by pressing a key.

In the calling phase the voice-control unit 2 has to be set to a choose-name mode, e.g. by a voice command "phone call" or by using the keys of the telecommunication terminal 1. When mounted on a car, it is also possible to bring a supplementary control option, external from the telecommunication terminal 1, e.g. close to the steering wheel of the car, wherein the activation of the choose-name mode is easy to implement, e.g. by an activation switch 14. In the following, the voice-controlled dialling of telephone number in accordance with a preferred embodiment of the invention is described with reference to the flow chart of Fig. 3.

After the voice-control unit 2 has recognised the given command as the activation command of the choose-name mode, the voice-control unit 2 moves to a choose-telephone-number mode (block 301). The voice-control unit 2 creates advantageously a sound signal to the loudspeaker 11 and/or a text message on a display means 13, which signal or message informs the user to pronounce the identifier (block 302). The user can pronounce the sub-identifiers of the identifier in any order, preferably by keeping a short pause between sub-identifiers to separate the sub-identifiers from each other. The voice-control unit 2 calculates the probability between the first stored identifier and the pronounced identifier (block 303). Subsequently, it is examined whether any other identifiers are stored into the memory (block 304). In case there remains any non-examined identifiers, probability is created for the next identifier (block 305). When probability has been created for every stored identifier, the highest calculated probability is searched. In case the probability calculated to one stored identifier is distinctively higher than that calculated to the rest of the identifiers, it can be assumed that the said identifier is the correct one (block 306), wherein choose-telephone-number mode can be started (block 307). In case the identifying of the identifier did not succeed, it is possible e.g. to move back to the block 302 and ask the user to repeat the identifier until the selection can be identified.

A complete identification is not always reached, wherein the voice-control unit 2 can inform the user and to ask the user to pronounce the identifier again, e.g. by moving back to block 302 in the flow chart of

- Fig. 3. The voice control unit 2 can also create e.g. a sound signal of those identifiers that according to the comparison made by the voice-recognition means 3 most resemble the pronounced identifier, wherein the user can select the correct identifier. In case none of the proposed
- 5 identifiers is correct, the user can repeat the identifier. Even if the voice-control unit 2 could recognize the given identifier, it is preferable to verify from the user that the selected identifier is correct. This can be performed for example in a manner that the user gives a dial command if the identifier is correct, or a re-recognition command if the identifier is
- 10 incorrect. The verifying can be advantageously performed also by an activation switch key 14. Yet another alternative for verifying is that the telecommunication terminal 1 will wait a predetermined time for the command of the user, and in case no command is coming, it presumes the selected telephone number to be correct and starts the dialling.
- 15 The telephone number is dialled according to the information stored to the telephone number memory in a manner known as such. The used memory can be memory of the telecommunication terminal 1 (not shown) or the random access memory 7 of the voice-control unit 2.
- 20 Also non-volatile random access memory (NVRAM) can be partially used as the random access memory 7 of the voice-control unit 2, wherein the information stored in the memory is preserved also without operating voltage.
- 25 The method according to the invention can be implemented e.g. in a manner that in the storing phase a separate model is formed of each pronounced identifier. In the following, it is assumed that N number of names, that is sub-identifiers: n_1, n_2, \dots, n_N , is related to the telephone number. For the recognition phase, a model structure is formed to the
- 30 telephone number, the structure including every possible sub-identifier composition, that is, 1 to N sub-identifiers in every possible order.

These sub-identifier compositions include

$$\sum_{i=1}^N i! \binom{N}{i} \text{ pcs.}$$

5

The voice-control unit 2 defines probability to all the sub-identifier compositions, and the sub-identifier composition which is given the highest probability is the final result of the recognition.

- 10 For example in the case $n_1 = \text{Williams}$, $n_2 = \text{Matthew}$ and $n_3 = \text{Herbert}$, the possible sub-identifier compositions are:

15

Williams, Matthew, Herbert, Williams Matthew, Matthew Williams, Williams Herbert, Herbert Williams, Matthew Herbert, Herbert Matthew, Williams Matthew Herbert, Williams Herbert Matthew, Matthew Williams Herbert, Matthew Herbert Williams, Herbert Williams Matthew, and Herbert Matthew Williams

- 20 Thus, there are altogether 15 possible sub-identifier compositions when the number of the sub-identifiers is three. Sub-identifier combinations are thus full combinations of sub-identifiers (consisting all the sub-identifiers) or partial combinations of sub-identifiers (consisting only a part of the sub-identifiers). Also partial combinations having only one
- 25 sub-identifier are possible when adapting the voice control according to the invention.

The following Table 1 shows the number of sub-identifier combinations as the function of sub-identifiers.

30

Number of sub-identifiers	Number of sub-identifier combinations
1	1
2	4
3	15

B As it can be seen in the Table 1, the number of sub-identifier combinations rises very quickly, it being as high as 64 when ^{the} ~~there~~ number of the sub-identifiers is four. The quantity of the memory and the calculation time required for storing the model structure can be diminished by means of the implementation alternative according to the preferred embodiment of the invention. In this alternative, separate sub-identifiers are recognised, these being independent from each other, from the group of all the pronounced words (word spotting). In this method, it looks as if the voice-control unit 2 is constantly waiting for a certain sub-identifier and it recognizes whether it is pronounced or not. In this case, the voice-control unit 2 produces several possible alternative names and a probability rank for them. According to these alternatives, the telephone number meant by the user can be concluded.

15 In this method, it does not make a difference how many words not included in the word list (the group of all the stored sub-identifiers) are used, which makes this method highly flexible in use.

20 In the teaching phase, the voice-control unit 2 transforms the pronounced sub-identifiers to a form appropriate for storing and compares each pronounced sub-identifier to ready-stored sub-identifiers. In case the pronounced sub-identifier had already been stored; e.g. the user has already stored the name "Matthew Taylor", the voice-control unit 2
 25 detects, when "Matthew" is being pronounced, that this had already been stored. In this case, the voice-control unit 2 forms a reference from the sub-identifier "Matthew" to the telephone number of Taylor and the telephone number of Williams. In this situation, in the recognition phase, after the sub-identifier "Matthew", the voice-control unit 2 has
 30 formed e.g. a list which includes both Matthew Taylor and Matthew Williams. Thus, the voice-control unit 2 knows to expect either Taylor or Williams, and after the user has pronounced the next sub-identifier, the voice-control unit 2 judges whether the identifier can be identified on basis of the given sub-identifiers or whether it should wait for a possible
 35 sub-identifier to come. This could be possible in such cases when the two sub-identifiers are identical and the third sub-identifier is different.

5

10

15

20

25

30

35

[illegible]

- 3